Barycenter and maximum likelihood ☆

Ruedi Flüge ✠, Ernst A. Ruh *

Department of Mathematics, University of Fribourg, CH-1700 Fribourg, Switzerland

Abstract

We refine recent existence and uniqueness results, for the barycenter of points at infinity of Hadamard manifolds, to measures on the sphere at infinity of symmetric spaces of non compact type and, more specifically, to measures concentrated on single orbits. The barycenter will be interpreted as the maximum likelihood estimate (MLE) of generalized Cauchy distributions on Furstenberg boundaries. As a spin-off, a new proof of the general Knight–Meyer characterization theorem will be given.

MSC: 53C35; 53C30; 14M15; 62H12

Keywords: Barycenter; Furstenberg boundaries; Bruhat Lemma; Cauchy distribution; Maximum likelihood estimate

1. Introduction

Ordinary Cauchy distributions on $\mathbb{R}^n$ may naturally be seen as the restriction, to an affine part in real projective space $\mathbb{P}^n$, of its standard Riemannian measure. By doing so hidden symmetries become apparent. And consequently a larger group, namely $\mathrm{SL}(n+1,\mathbb{R})$, acts on these measures leaving the type, i.e. the images under affine transformations, invariant. Moreover this fact characterizes the Cauchy type. This is the content of the original *Knight–Meyer characterization theorem*. In Section 4 we give a new proof of it in the general setting:

Definition 1. Let G be a semi simple Lie group of non compact type and Q a parabolic subgroup. The (generalized) flag manifold G/Q is also called a *Furstenberg boundary*. The family of K-invariant probability measures on G/Q, $K \subset G$ maximal compact, is said to be the (generalized) *Cauchy type*.

Let K be maximal compact and denote by $\mu_K \in \mathcal{M}^1(G/Q)$ the K-invariant member in the set of probability measures on G/Q. If a Furstenberg boundary G/Q is *faithful*, the map $\phi: G/K \to \mathcal{M}^1(G/Q)$, $gK \mapsto g\mu_K$ is injective—see below. This happens for example if G is simple. The family of distributions is then parametrized by the symmetric space (of non compact type) $M = G/K$. Moreover the stabilizer subgroups G_ξ of points at infinity

☆ The authors were supported by the Swiss National Science Foundation Grants 20-67619.02, 200020-10501/1.
* Corresponding author.
E-mail address: ernst.ruh@unifr.ch (E.A. Ruh).
✠ Ruedi Flüge died December 24, 2005.

$\xi \in M(\infty)$ are parabolic subgroups, thus the G-orbits in the sphere at infinity $M(\infty)$ are Furstenberg boundaries. We therefore are in the following situation: *the sample space is part of the boundary of the parameter space*. That is the point where Riemannian geometry comes into the picture: The notion of barycenter, whose existence and uniqueness is discussed in Sections 2 and 3, coincides with the maximum likelihood estimate (MLE) of generalized Cauchy distributions, see Section 5. This observation yields a very conceptional treatment of MLE by geometric reasoning.

Finally we like to mention two important ideas on which our paper is based: convexity and non-positive curvature, as well as the (generalized) Bruhat double coset lemma which is fundamental to the study of parabolic subgroups.

2. The barycenter

Let M be a Hadamard manifold and denote by $M(\infty)$ the sphere at infinity. Any point at infinity $\xi \in M(\infty)$ determines a unit vector field X_ξ on M:

$$X_\xi = \text{unit vector at } p \text{ pointing to } \xi.$$

These vector fields admit potentials, namely minus one times the *Busemann functions* $b_\xi(p)$. These are known to be convex (see [9]).

If a Borel probability measure μ is given on $M(\infty)$, we form the following integrals:

$$X_\mu(p) = \int_{M(\infty)} X_\xi(p)\mu(d\xi) \quad \text{and} \quad b_\mu(p) = \int_{M(\infty)} b_\xi(p)\mu(d\xi).$$

Note that b_μ is convex as well.

Definition 2. A point $p \in M$ is called a *barycenter* of μ if it satisfies the following equivalent conditions:

(1) $X_\mu(p) = 0$,
(2) p is a minimum of b_μ.

This notion of barycenter of measures at infinity has already been studied and fruitfully used for example in [5,7]. The set of barycenters of a measure μ is a convex subset of M. Recall that on $M(\infty)$ a well known metric is given by the *angle metric*:

$$\angle(\xi,\eta) = \sup_{p\in M} \angle_p\big(X_\xi(p), X_\eta(p)\big) = \lim_{t\to\infty} \angle_{\gamma(t)}\big(X_\xi\big(\gamma(t)\big), X_\eta\big(\gamma(t)\big)\big).$$

Where γ is a geodesic ray representing ξ, we write $\xi = \gamma(\infty)$. The last equality is Proposition 3.1.3 in [9]. The link to the asymptotic behavior of b_μ is this: Let γ be a unit speed geodesic in M and consider the convex function $f(t) = b_\mu(\gamma(t))$. The derivative satisfies

$$f'(t) = -\int_{M(\infty)} \big\langle X_\xi\big(\gamma(t)\big), \dot\gamma(t)\big\rangle \mu(d\xi) = -\int_{M(\infty)} \cos\angle\big(X_\xi\big(\gamma(t)\big), X_\eta\big(\gamma(t)\big)\big)\mu(d\xi)$$

by Lebesgue's dominated convergence, where $\eta = \gamma(\infty)$. The limit as $t \to \infty$ is

$$C(\mu,\eta) = \int_{M(\infty)} \cos\angle(\xi,\eta)\mu(d\xi). \tag{1}$$

Its sign tells you whether f converges to a finite value or to $\pm\infty$ as $t \to \infty$. This was observed in [1] to prove the following

Theorem 3. *Let μ be a Borel probability measure on $M(\infty)$, where M is a Hadamard manifold with either strictly negative curvature or with an analytic metric. Then there exists a unique barycenter of μ if and only if μ satisfies the* obtuse angle condition*, that is if*

$$C(\mu,\xi) < 0 \quad \textit{for all } \xi \in M(\infty).$$

More details as well as the case of general Hadamard manifolds may be found in [1]. In the case of strictly negative curvature, which includes rank one symmetric spaces, the criterion becomes quite simple. Since then any pair of distinct ideal points can be joined by a unique geodesic, one gets

$$C(\mu,\xi) = \mu\big(\{\xi\}\big) - \mu\big(M(\infty)\setminus\{\xi\}\big) = 2\mu\big(\{\xi\}\big) - 1$$

and a unique barycenter exists if and only if $\mu(\{\xi\}) < \frac{1}{2}$ for all $\xi \in M(\infty)$.

3. Bruhat Lemma

We turn to the case of symmetric Hadamard manifolds. Since the behavior of the barycenter of measures at infinity is rather pathological for euclidean spaces, it is natural to consider in the sequel symmetric spaces of *non compact type*. That is $M = G/K$ where G is a semi simple (real) Lie group with finite center and K a maximal compact subgroup. We need some facts of their rich structure. First the sphere at infinity partitions into equivalence classes given by

$$\mathcal{C}(\xi) = \big\{\eta \in M(\infty) \mid G_\xi = G_\eta\big\}.$$

These are partially ordered by

$$\mathcal{C}(\xi) \leqslant \mathcal{C}(\eta) \iff G_\xi \supset G_\eta.$$

The maximal ones are called *Weyl chambers* and the others (Weyl) *faces*. Denote by $\bar{\mathcal{C}}(\xi) = \{\eta \in M(\infty) \mid \mathcal{C}(\eta) \leqslant \mathcal{C}(\xi)\}$, the set of points lying in the faces of the down set of $\mathcal{C}(\xi)$. It is the closure of $\mathcal{C}(\xi)$ in the cone topology, and is called the closed chamber or face, see [9]. A fixed closed Weyl chamber, or the set of standard parabolic subgroups, may be identified—as a poset—with not co-empty subsets of $\{1, 2, \ldots, r\}$ for some r known as the rank of M. This allows a consistent labeling of all chambers and faces. A flat at infinity is $F(\infty) = \{\gamma(\infty) \mid \gamma \subset F \text{ for some fixed maximal flat } F \subset M\} \subset M(\infty)$. We have the following

Proposition 4. *Let $\xi, \eta \in M(\infty)$, $p_0 \in M$ and $a, b \in \mathbb{R}_{>0}$. Let $\zeta = \gamma(\infty)$, where γ is determined by $\gamma(0) = p_0$ and $\dot\gamma(0) = \frac{1}{\|aX_\xi + bX_\eta\|}(aX_\xi + bX_\eta)(p_0)$. Then the following are equivalent*:

(1) $\xi, \eta \in \bar{\mathcal{C}}(\zeta)$, *i.e.* $\mathcal{C}(\xi) \leqslant \mathcal{C}(\zeta)$ *and* $\mathcal{C}(\eta) \leqslant \mathcal{C}(\zeta)$;
(2) *if* $\xi, \eta \in F(\infty)$ *then* $\zeta \in F(\infty)$;
(3) $aX_\xi + bX_\eta = \|aX_\xi + bX_\eta\| X_\zeta$;
(4) $ab_\xi + bb_\eta = \|aX_\xi + bX_\eta\| b_\zeta$;
(5) $\angle_p(X_\xi, X_\eta) \equiv \mathrm{const}$ *for all* $p \in M$.

Proof. The equivalence (1) $\iff$ (2) follows immediately from Proposition 3.6.26 in [9]. Next (1) $\Longrightarrow$ (3): since (3) holds at p_0 it holds everywhere in M as G_ζ is transitive on M. (3) $\iff$ (4) is clear. To prove (4) $\iff$ (5) we note that a function f on M is a Busemann function if and only if following properties are satisfied (see [4, p. 24]):

(1) f is a convex C^1-function,
(2) $\|\operatorname{grad} f\| \equiv 1$ on M.

Hence Busemann functions 'add' as stated above if and only if the angle between their gradients remains constant. Finally to prove (5) $\Longrightarrow$ (2) note that (5) implies $\angle(\xi, \eta) = \angle_p(X_\xi, X_\eta) < \pi$ for all $p \in M$—since there are no Euclidean factors—and that happens if and only if $X_\xi(p)$ and $X_\eta(p)$ span a flat sector (see [9, Proposition 3.1.2]) that contains $\frac{1}{\|aX_\xi + bX_\eta\|}(aX_\xi + bX_\eta)(p)$ and (2) follows. □

This proposition can be interpreted in the following way: fix $\xi \in M(\infty)$ maximal (regular) then the pointed Weyl chamber at $p \in M$ asymptotic to $\mathcal{C}(\xi)$ is given by $\sum_{i=1}^r a_i X_{\xi_i}(p)$ where $a_i > 0$ and $\xi_i \in \bar{\mathcal{C}}(\xi)$ are minimal. Now we can sharpen the criterion in Theorem 3.

Corollary 5. *Let μ be a probability measure on $M(\infty)$ where M is a symmetric space of non compact type. Then a unique barycenter exists if and only if*

$$C(\mu,\xi)<0 \quad \text{for all minimal } \xi \in M(\infty).$$

Proof. Let $\eta \in M(\infty)$ then X_η can be expressed as $X_\eta = \sum_{i=1}^{n} a_i X_{\xi_i}$ with $\xi_i \in \bar{\mathcal{C}}(\eta)$ minimal and $a_i \geqslant 0$ by the proposition above. Thus

$$C(\mu,\eta)=\lim_{t\to\infty}\int_{M(\infty)} \langle X_\xi\big(\gamma(t)\big), X_\eta\big(\gamma(t)\big)\rangle \mu(d\xi)=\sum_{i=1}^{n} a_i C(\mu,\xi_i),$$

where $\gamma(\infty)=\eta$, is negative if the $C(\mu,\xi_i)$ are. Necessity is clear. □

It follows from basic properties of root systems that $\angle_p(X_\xi(p), X_\eta(p)) \leqslant \frac{\pi}{2}$ if $\xi,\eta\in\mathcal{C}$ and equality can occur only if M is reducible. To study the semi simple case, it is sensible to consider only *faithful* Furstenberg boundaries:

Definition 6. Let $M = M_1\times M_2\times\cdots\times M_l$ be the decomposition of M into its irreducible factors and $\gamma(t)=(\gamma_1(t),\gamma_2(t),\ldots,\gamma_l(t))$ a geodesic, where $\gamma_i(t)\in M_i$ for all $t\in\mathbb{R}$ and $1\leqslant i\leqslant l$. The Furstenberg boundary G/G_ξ with $\xi=\gamma(\infty)$ is called faithful if $\dot\gamma_i(t)\neq 0$ for all $1\leqslant i\leqslant l$.

Closed Weyl chambers are fundamental domains for the action of G on $M(\infty)$:

$$M(\infty)=\coprod_{\xi\in\bar{\mathcal{C}}} G\xi. \tag{2}$$

Furthermore $G\xi\cong G/G_\xi$ and the stabilizer G_ξ is a parabolic subgroup of G. Let W be the *Weyl group* of G. The (generalized) *Bruhat Lemma* or *Bruhat decomposition* states that

$$G=\coprod_{w\in W_\eta\backslash W/W_\xi} G_\eta w G_\xi \quad \text{or} \quad G/G_\xi=\coprod_{w\in W_\eta\backslash W/W_\xi} G_\eta[w], \tag{3}$$

where $\xi,\eta\in\bar{\mathcal{C}}$ and $[\,.\,]$ denotes cosets. This is a cellular decomposition of G or G/G_ξ respectively, if η is regular that is to say that the stabilizer subgroup W_η is trivial. There is exactly one orbit of maximal dimension—called large one—namely $G_\eta w^* G_\xi$ or $G_\eta[w^*]$ respectively where $w^*\in W$ is the element that sends the Weyl chamber $\mathcal{C}$ to its opposite $-\mathcal{C}$, see [17, pp. 49 and 76]. Let $\mu\in\mathcal{M}^1(M(\infty))$ be a probability measure that is concentrated on a single G-orbit. The function

$$\begin{aligned} &G/G_\xi \longrightarrow \mathbb{R}\\ &g\xi \mapsto \angle(\eta, g\xi)=\angle\big(g^{-1}\eta,\xi\big), \end{aligned}$$

where $\xi,\eta\in\bar{\mathcal{C}}$ as above pushes down to $G_\eta\backslash G/G_\xi \simeq W_\eta\backslash W/W_\xi$ and hence is a simple function with level sets the G_η-orbits. Plugging that into (1) we get

$$\begin{aligned} C(\mu,\eta) &= \int_{M(\infty)} \cos\angle(\eta,\xi)\mu(d\xi) = \int_{G/G_\xi} \cos\angle(\eta, g\xi)\mu(dg\xi)\\ &= \sum_{w\in W_\eta\backslash W/W_\xi} \cos\angle(\eta, w\xi)\mu\big(G_\eta[w]\big). \end{aligned}$$

Given $\eta\in M(\infty)$, then for any $\xi\in M(\infty)$ there is a $g\in G$ such that there is a closed Weyl chamber $\bar{\mathcal{C}}$ with $\eta, g\xi\in\bar{\mathcal{C}}$, by the fundamental domain property of closed Weyl chambers, i.e. (2). These remarks yield an other corollary to Theorem 3:

Corollary 7. *Let μ be a Borel probability measure on $M(\infty)$ that is concentrated on one G-orbit $G\xi_0$. Then a unique barycenter exists if and only if*

$$\sum_{w \in W_\eta \backslash W / W_\xi} \cos\angle(\eta, w\xi)\mu\big(G_\eta[w]\big) < 0 \quad \text{for all (minimal) } \eta \in M(\infty),$$

where $\xi = g\xi_0$ and η lie in some common closed Weyl chamber $\bar{\mathcal{C}}$. In particular, if G/G_ξ is faithful, a unique barycenter exists for probability measures that are absolutely continuous w.r.t. a quasi invariant measure.

Proof. Only the last statement needs further care: the G_η-orbits $G_\eta[w]$ are either submanifolds of positive codimension and consequently $\mu(G_\eta[w]) = 0$ or the large cell. In the latter case one has $\angle(\eta, w^*\xi) > \frac{\pi}{2}$ since η and $w^*\xi$ lie in opposite closed Weyl chambers and G/G_ξ is assumed faithful. □

Recall that $M(\infty)$ may be identified with a sphere $S_p \subset T_pM$ in some tangent space.

Corollary 8. *Let μ be a probability measure on $M(\infty)$ that is absolutely continuous w.r.t. the standard measure of a unit sphere S_p at some $p \in M$. Then it has a unique barycenter.*

Proof. For $\eta \in M(\infty)$ choose ζ regular such that $\eta \in \overline{\mathcal{C}(\zeta)}$ and let $\mathcal{C}(\xi) \leqslant \mathcal{C}(\zeta)$ be the faces. Then (2) and (3) read as

$$M(\infty) = \coprod_{\mathcal{C}(\xi) \leqslant \mathcal{C}(\zeta)} G\mathcal{C}(\xi) = \coprod_{\mathcal{C}(\xi) \leqslant \mathcal{C}(\zeta)} \coprod_{w \in W_\eta \backslash W / W_\xi} G_\eta \mathcal{C}(w\xi)$$

and all subsets in that partition except for $G_\eta\mathcal{C}(w^*\zeta)$ have positive codimension whence

$$C(\mu, \eta) = \int_{M(\infty)} \cos\angle(\eta, \xi)\mu(d\xi) = \int_{G_\eta\mathcal{C}(w^*\zeta)} \cos\angle(\eta, \xi)\mu(d\xi) < 0$$

as above. □

4. Knight–Meyer characterization of the generalized Cauchy type

In [15,16] the dynamic characterization of Cauchy distributions was discovered. This was generalized to Furstenberg boundaries in [8] and the final version was settled in [6]. It says:

Theorem 9. *Let G be a semi simple Lie group of non compact type and μ a Borel probability measure on a Furstenberg boundary G/Q. Then $Q\mu = G\mu$ if and only if μ is K-invariant for some maximal compact subgroup $K \subset G$.*

It is enough to show the result for faithful boundaries. The proof is easy once one knows that the stabilizer subgroup G_μ of μ is compact: the hypothesis on μ implies $G_\mu \cap Qg \neq \emptyset$ which is the same as $G_\mu \cap gQ \neq \emptyset$ for all $g \in G$. Thus G_μ acts transitively on G/Q and μ is the normalized Haar measure of some maximal compact $K \supset G_\mu$. Necessity follows from the decomposition of $G = KQ = QK$.

We establish compactness of G_μ: consider the large Q-orbit $Q[w^*] \subset G/Q$ and its complement $\partial Q[w^*]$. Since $G\mu = Q\mu$ it follows that $\mu(g(Q[w^*])) = \mu(Q[w^*])$ and $\mu(g(\partial Q[w^*])) = \mu(\partial Q[w^*])$ for all $g \in G$. Let λ be a Haar measure on G and denote by $\phi_{\partial Q[w^*]}$ the characteristic function of $\partial Q[w^*]$. Then $\lambda(\partial Q w^* Q) = 0$ in G since it is a finite union of submanifolds with positive codimension. Thus for all $h \in G$

$$0 = \lambda(\partial Q w^* Q) = \int_G \phi_{\partial Q[w^*]}\big(g[h]\big)\lambda(dg) = \int_{G/Q} \left(\int_G \phi_{\partial Q[w^*]}\big(g[h]\big)\lambda(dg)\right)\mu\big(d[h]\big)$$

$$= \int_G \left(\int_{G/Q} \phi_{\partial Q[w^*]}\big(g[h]\big)\mu\big(d[h]\big)\right)\lambda(dg) = \int_G \mu\big(g^{-1}\partial Q[w^*]\big)\lambda(dg)$$

by Fubini's theorem. Hence $\mu(g\partial Q[w^*]) = 0$ a.e. and consequently everywhere since it is a constant function of g. That implies together with Corollary 7 that the measures to be characterized have a unique barycenter $p_0 \in M \cong G/K$. On the other hand the obvious equivariance of the barycenter, that is

$$p \in M \text{ is a barycenter of } \mu \iff gp \in M \text{ is a barycenter of } g\mu,$$

implies $G_\mu \subseteq G_{p_0}$, which is a maximal compact subgroup of G as claimed in Theorem 9.

In [2] the results of [3] have been announced and used to prove the original characterization theorem along the lines as above. This completes also the discussion on faithfulness:

Corollary 10. *A Furstenberg boundary G/Q is faithful if and only if $G_{\mu_K} = K$ where μ_K is the K-invariant probability on G/Q for $K \subset G$ maximal compact. That is if and only if the map*

$$\begin{aligned} \phi: G/K &\to \mathcal{M}^1(G/Q) \\ gK &\mapsto g\mu_K \end{aligned}$$

is injective.

Remark 11. (*Compare Remark* 3 *in* [6]) In all papers, we are aware of, on this characterization theorem except for [15] and [6], there is the additional hypothesis on μ that it does not charge thin cells, called 'Condition C'. Recalling Furstenberg's original definition of G-boundaries, that is compact homogeneous G-spaces F with the property that for any $\nu \in \mathcal{M}^1(F)$ there is a sequence $\{g_n\} \subset G$ such that $g_n\nu$ converges weakly to a point measure, see [10], yields an alternative proof of the fact that $G\mu = Q\mu$ implies $\mu(\partial Q[w^*]) = 0$: Take $\{g_n\}$ with $g_n\mu \to \delta_{[e]}$. Since $Q\mu = G\mu$ we can assume $q_n = g_n \in Q$. Let $0 \leqslant f \in \mathcal{C}(G/Q)$ with $f([e]) = 0$ and $f \equiv 1$ in a neighborhood of the complement of the large orbit $\partial Q[w^*]$, then

$$\mu\big(\partial Q[w^*]\big) = \int\limits_{\partial Q[w^*]} \mu \leqslant \int\limits_{G/Q} f(q_n x)\, d\mu(x) = q_n\mu(f) \overset{n\to\infty}{\longrightarrow} f\big([e]\big) = 0$$

as Q leaves the large orbit invariant.

5. Maximum likelihood estimate of generalized Cauchy distributions

Estimation of course is about finding the suitable parameter in certain statistical model of a given sample: Let $(X, \mathcal{B}, p_\theta)_{\theta\in\Theta}$ be a (parametric) statistical model, $\{x_i\}_{i=1\ldots N}$ a (random) sample and $\mu = \frac{1}{N}\sum \delta_{x_i}$ the corresponding empirical measure. Assume continuous (relative) density functions $\frac{dp_\theta}{dp_{\theta_0}}$. The *likelihood function* of that sample is defined by

$$L(\mu,\theta) = \prod_{i=1}^{N} \frac{dp_\theta}{dp_{\theta_0}}(x_i).$$

The parameter that maximizes that function for a given observation is called the *maximum likelihood estimate* and is denoted by MLE(μ), if it exists and is unique. The *log likelihood function* is given by

$$\ell(\mu,\theta) = \log L(\mu,\theta) = \sum_{i=1}^{N} \log \frac{dp_\theta}{dp_{\theta_0}}(x_i).$$

It has the same maxima as $L(\mu,\theta)$ and can be defined for an arbitrary probability measure μ:

$$\ell(\mu,\theta) = \int\limits_X \log \frac{dp_\theta}{dp_{\theta_0}}(x)\mu(dx).$$

This expression suggests how to link the barycenter to the MLE of generalized Cauchy laws on Furstenberg boundaries: Under what conditions is the logarithm of the density function of our Cauchy distributions proportional to the corresponding Busemann function, that is

$$\log \frac{dg\mu_K}{d\mu_K}(h\xi) = cb_{h\xi}(gp) \quad \text{for all } g, h \in G,$$

where $p \in M$, $K = G_p$ and μ_K is the K-invariant probability on G/G_ξ? Roughly the answer is that one has to choose ξ in the 'middle' of the corresponding chamber or face.

Theorem 12. *Let G/Q be a Furstenberg boundary, μ a probability measure of Cauchy type, $p \in M$, $K \subset G$ and μ_K as above. Then there is a point at infinity $\xi \in M(\infty)$ and $c \in \mathbb{R}$ such that*

(1) $G/Q \cong G\xi \subset M(\infty)$,
(2) $\frac{dg\mu_K}{d\mu_K}(h\xi) = \mathrm{e}^{cb_{h\xi}(gp)}$ *for all* $g, h \in G$.

With this choice, the barycenter and MLE of a probability measure on G/Q coincide if they exist.

Although the result follows from inspection of the relevant formulæ, compare the discussion below, we give an elementary

Proof. (*Compare* [11] *and* [18] *Section* 4.2) A (Borel) function $\sigma : G \times S \to \mathbb{R}$, where G is a (topological) group and S a G-space, is called a (Borel) cocycle if

$$\sigma(gh, s) = \sigma(g, hs) + \sigma(h, s) \quad \text{for all } g, h \in G \text{ and } s \in S.$$

It follows that the function $\phi_{\sigma,s} : G_s \to \mathbb{R}$, defined by $\phi_{\sigma,s}(g) = \sigma(g, s)$, is a homomorphism of the stabilizer subgroup G_s to $\mathbb{R}$. Let $K \subset G$ be a subgroup. Then a cocycle is called K-invariant if $\sigma(k, s) = 0$ for any $k \in K$ and $s \in S$. If moreover S is homogeneous and $K \subset G$ is transitive on $S = G/G_{s_0}$, i.e. $G = KG_{s_0} = G_{s_0}K$, then a K-invariant cocycle σ is uniquely determined by ϕ_{σ,s_0}:

$$\sigma(g, ks_0) = \sigma(gk, s_0) - \sigma(k, s_0) = \sigma(k'q, s_0) - 0 = \sigma(k', qs_0) + \sigma(q, s_0) = \phi_{\sigma,s_0}(q),$$

where $gk = k'q$ with $q \in G_{s_0}$ and $k, k' \in K$. In particular, if G is our semi simple Lie group, $K = G_p$ for a fixed $p \in M$ and $\xi \in M(\infty)$ then the functions

$$\begin{aligned} \sigma_1 : G \times G/G_\xi &\to \mathbb{R} \\ (g, h\xi) &\mapsto \log \frac{dg^{-1}\mu_K}{d\mu_K}(h\xi) \end{aligned} \qquad \text{and} \qquad \begin{aligned} \sigma_2 : G \times M(\infty) &\to \mathbb{R} \\ (g, \xi) &\mapsto b_\xi\big(g^{-1}p\big) \end{aligned}$$

where the Busemann functions are normalized such that $b_\xi(p) = 0$ for all $\xi \in M(\infty)$, are K-invariant cocycles. Indeed, this follows from the chain rule for the Radon–Nikodym derivative for σ_1 and from the definition of the Busemann cocycle:

$$b_\xi(q) = \lim_{t\to\infty} d\big(q, \gamma(t)\big) - d\big(p, \gamma(t)\big),$$

where $\gamma(\infty) = \xi$ whence

$$\begin{aligned} b_\xi\big((gh)^{-1}p\big) &= \lim_{t\to\infty} d\big(h^{-1}g^{-1}p, \gamma(t)\big) - d\big(p, \gamma(t)\big) \\ &= \lim_{t\to\infty} d\big(g^{-1}p, h\gamma(t)\big) - d\big(p, h\gamma(t)\big) + d\big(h^{-1}p, \gamma(t)\big) - d\big(p, \gamma(t)\big) \\ &= b_{h\xi}\big(g^{-1}p\big) + b_\xi\big(h^{-1}p\big) = \sigma_2(g, h\xi) + \sigma_2(h, \xi). \end{aligned}$$

On the other hand any parabolic subgroup can be written as $Q = K_Q AN$, where $K_Q = K \cap Q$ and $KAN = G$ is a corresponding Iwasawa decomposition of G. Furthermore one has $N \subset [Q, Q]$—as $[AN, AN] = N$—whence any real homomorphism is already determined by its restriction to A. The image of A in $M = G/K$ is the pointed flat F

that we identify—as an abelian group—with a subspace $F \subset T_pM$ and any real Borel homomorphism is of the form

$$F \ni Y \mapsto \langle Y, X\rangle = \|X\|\left\langle Y, \frac{X}{\|X\|}\right\rangle \quad \text{for some } X \in F.$$

Which in turn is a multiple of a Busemann function of F and, since F is totally geodesic, it is the restriction of b_ξ on M with $\xi = \gamma(\infty) \in F(\infty)$ where $\gamma(0) = p$ and $\dot\gamma(0) = \frac{X}{\|X\|}(p)$. Thus

$$\log \frac{dg\mu_K}{d\mu_K}(hQ) = cb_{h\xi}(gp).$$

It is clear that Q stabilizes ξ. □

To make more precise what is meant by the middle of the chamber or face and to point out the connection to harmonic analysis, we need some notation: Let $\xi \in G/K(\infty)$ as in Theorem 12, take the Iwasawa decomposition of $G = KAN$ such that $G_\xi = (G_\xi \cap K)AN$ and denote by $H(g) \in \mathfrak{a}$ the logarithm of the A-component of $g = kan$, i.e. $H(g) = \log(a)$. Denote further $\rho_\xi = \sum_{\alpha(X_\xi(p))>0} m_\alpha \alpha$ where $m_\alpha = \dim \mathfrak{g}_\alpha$ is the multiplicity of the root α. Now

$$\frac{dg\mu_K}{d\mu_K}(k\xi) = \operatorname{const} \mathrm{e}^{-\rho_\xi(H(g^{-1}k))} \quad k \in K \text{ and } g \in G.$$

In other words $\langle X_\xi(p), Y\rangle = c\rho_\xi(Y)$ for all $Y \in \mathfrak{a}$; the ‘middle’ however is intuitive only if $\mathfrak{g}$ is a normal real form. And one recognizes a typical instance of the integrand in Harish-Chandra’s integral formula for elementary spherical functions. It is a well known fact that the relative densities of K-invariant measure on flag manifolds are of this form, see [12–14] or your favored reference to the subject. Alternatively this formula can be derived along similar lines of reasoning as in the proof of Theorem 12 by observing that $(g, k) \mapsto \lambda(H(g^{-1}k))$ defines a K-invariant cocycle for any $\lambda \in \mathfrak{a}^*$. Analogous remarks apply to Busemann functions.

5.1. MLE for empirical measures

In application empirical measures are of prominent importance.

Proposition 13. *Let $G\xi$, where $\xi \in M(\infty)$ is chosen according to Theorem 12, be a Furstenberg boundary then there is an integer $N \in \mathbb{N}$ such that MLE exists for almost all—w.r.t. the a quasi invariant measure—samples of size bigger or equal to N.*

Proof. Let $\bar{\mathcal{C}}$ be an arbitrary closed Weyl chamber and $\xi_i \in \bar{\mathcal{C}}$ be minimal for $i = 1, \ldots, r$, where r is the rank of M and $g_0\xi = \xi_0 \in \bar{\mathcal{C}}$. Choose a system of representatives $\{w^i_j\}_{j=0,\ldots,n_i}$ of $W_{\xi_i}\backslash W/W_{\xi_0}$, with $w^i_0 = w^*$ for all i, and denote $c^i_j = \cos\angle(\xi_i, w_j\xi_0)$. Since the G_ξ-orbits $G_{\xi_i}[w^i_j]$ are real algebraic varieties, of positive codimension if $j \neq 0$, there is an $N^i_j \in \mathbb{N}$ such that the set $\{x = (x_1, \ldots, x_n) \in (G\xi)^n$ s. t. there is no G_η-orbit B with $x_k \in B$ for $k = 1, \ldots, n\}$ is Zariski open dense and hence has full measure if $n \geqslant N^i_j$. Put

$$N = \min\left\{n \in \mathbb{N} \mid n > \frac{|c^i_j - c^i_0|N^i_j n_i}{|c^i_0|} \text{ for all } i = 1, \ldots, r \text{ and } j = 1, \ldots, n_i\right\}.$$

Let $\mu = \frac{1}{L}\sum_{k=1}^{L}\delta_{x_k}$ be the empirical measure of a sample $(x_1, \ldots, x_L) \in (G/G\xi)^L$ with $L \geqslant N$. Then almost surely

$$\begin{aligned} C(\mu, \xi_i) &= \sum_{j=0}^{n_i} c^i_j \mu\big(G_{\xi_i}[w_j]\big) = c^i_0 + \sum_{j=1}^{n_i}\big(c^i_j - c^i_0\big)\mu\big(G_{\xi_i}[w^i_j]\big) \leqslant c^i_0 + \sum_{j=1}^{n_i}\big|c^i_j - c^i_0\big|\frac{N^i_j}{L} \\ &\leqslant c^i_0 + n_i \max\left\{\big|c^i_j - c^i_0\big|\frac{N^i_j}{L},\ j = 1, \ldots n_i\right\} < 0 \end{aligned}$$

since $c^i_0 < 0$ and by the choice of N. □

Let us finally inspect the classical case of ordinary Cauchy laws. The choice of a Weyl chamber in $\mathrm{SL}(n+1)/\mathrm{SO}(n+1)(\infty)$ corresponds to that of a full flag $U_1 \subsetneq U_2 \subsetneq \cdots \subsetneq U_n \subsetneq \mathbb{R}^{n+1}$ in $\mathbb{R}^{n+1}$, i.e. a maximal chain of subspaces. The lattice of standard parabolic subgroups of $\mathrm{SL}(n+1,\mathbb{R})$ is then given by stabilizers of flags in $\mathbb{R}^{n+1}$ of the form

$$U_{i_1} \subsetneq U_{i_2} \subsetneq \cdots \subsetneq U_{i_k} \subsetneq \mathbb{R}^{n+1}.$$

In particular the maximal parabolic subgroups (corresponding to minimal faces) are exactly the stabilizer subgroups of the proper subspaces $U \subsetneq \mathbb{R}^{n+1}$. Hence projective space is indeed a Furstenberg boundary:

$$\mathbb{P}^n = \mathrm{SL}(n+1,\mathbb{R})/\,\mathrm{SL}(n+1,\mathbb{R})_L \text{ with } L \subset \mathbb{R}^{n+1} \text{ and } \dim L = 1.$$

The restriction to an affine part of the $\mathrm{SO}(n+1)$-invariant probability on $\mathbb{P}^n$ yields, in fact, an ordinary multivariate Cauchy distribution. Let $U \subsetneq \mathbb{R}^{n+1}$ be a proper subspace and denote by $\mathbb{P}(U)$ the projective subspace induced by U. Choose $x, y \in \mathbb{P}^n$ with $x \in \mathbb{P}(U)$ and $y \notin \mathbb{P}(U)$. Then the Bruhat decomposition, see (3), reads as

$$\mathbb{P}^n = \mathrm{SL}(n+1,\mathbb{R})_U x \cup \mathrm{SL}(n+1,\mathbb{R})_U y = \mathbb{P}(U) \cup \mathbb{P}(U)^c,$$

where $\mathbb{P}(U)^c = \mathbb{P}^n \setminus \mathbb{P}(U)$ is the complement. By Corollary 7, a probability measure μ on $\mathbb{P}^n$ has a unique barycenter if and only if

$$c_0^k \mu\big(\mathbb{P}(U)^c\big) + c_1^k \mu\big(\mathbb{P}(U)\big) = c_0^k\big(1 - \mu\big(\mathbb{P}(U)\big)\big) + c_1^k \mu\big(\mathbb{P}(U)\big) < 0$$

or

$$\mu\big(\mathbb{P}(U)\big) < \frac{-c_0^k}{c_1^k - c_0^k}$$

for any subspace U where $k = \dim U$. Observe that minimal faces are 0-dimensional and the choice of the ‘middle’ becomes vacuous. Let $\mathfrak{a} \subset \mathfrak{sl}(n+1,\mathbb{R})$ be the abelian subalgebra consisting of diagonal elements. It is the tangent space of a typical flat. Denote $X_k \in \mathfrak{a}$ the element with the first k (diagonal) entries equal to $n+1-k$ and the remaining ones equal to $-k$. One finds then

$$c_1^k = \frac{\langle X_1, X_k\rangle}{\|X_1\|\|X_k\|} \quad \text{and} \quad c_0^k = \frac{\langle X_1, w^* X_k\rangle}{\|X_1\|\|X_k\|}$$

and

$$\langle X_1, X_k\rangle = n(n+1-k) - (k-1)(n+1-k) + (n+1-k)k = (n+1)^2 - k(n+1),$$
$$\langle w^* X_1, X_k\rangle = -k(n+1-k) + (n-k)k - kn = -k(n+1),$$

since w^* acts on $\mathfrak{a}$ by reversing the order of the diagonal entries. Thus

$$\frac{-c_0^k}{c_1^k - c_0^k} = \frac{k}{n+1}$$

and we get:

Proposition 14. *A probability measure $\mu \in \mathcal{M}^1(\mathbb{P}^n)$ on real projective space $\mathbb{P}^n$ has a unique barycenter if and only if*

$$\mu\big(\mathbb{P}(U)\big) < \frac{\dim U}{n+1}$$

for all proper subspaces $U \subsetneq \mathbb{R}^{n+1}$. In particular the MLE of an empirical measure exist almost surely if and only if the sample has size bigger than $n+1$.

This last proposition follows from Theorem 1 in [3], whose success motivated the investigation of the general pattern behind that result.

References

[1] C. Auderset, Barycenter of points at infinity, Preprint, Fribourg 2004.
[2] C. Auderset, R. Flüge, E. Ruh, Characterization of angular Gaussian distributions, in: Proceedings of the Hawaii International Conference on Statistics and Related Fields 2003, Hawaii 2003, ISSN #: 1539-7211.
[3] C. Auderset, C. Mazza, E. Ruh, Angular Gaussian and Cauchy estimation, J. Multivariate Analysis 93 (2005) 180–197.
[4] W. Ballmann, M. Gromov, V. Schroeder, Manifolds of Nonpositive Curvature, Progress in Mathematics, vol. 61, Birkhäuser.
[5] G. Besson, G. Courtois, S. Gallot, Entropies et rigidités des espaces localemnat symétriques de courbure strictement négative, GAFA 5 (5) (1995) 731–799.
[6] S.G. Dani, A characterization of Cauchy-type distributions on boundaries of semisimple groups, J. Theor. Prob. 4 (3) (1991) 229–625.
[7] A. Douady, C.J. Earle, Conformally natural extension of homeomorphisms of the circle, Acta Mathematica 157 (1986) 23–48.
[8] J.-L. Dunau, H. Sénateur, Characterization of the type of some generalizations of the Cauchy distribution, in: H. Heyer (Ed.), Probability Measures in Groups IX, Oberwolfach, in: Lecture Notes in Mathematics, vol. 1379, Springer-Verlag, 1988, pp. 64–74.
[9] P.B. Eberlein, Geometry of Nonpositively Curved Manifolds, The University of Chicago Press, 1996.
[10] H. Furstenberg, A Poisson formula for semi-simple Lie groups, Ann. Math. 77 (1963) 335–386.
[11] H. Furstenberg, I. Tzkoni, Spherical functions and integral geometry, Israel J. Math. 10 (1971) 327–338.
[12] R. Gangolli, V.S. Varadarajan, Harmonic Analysis of Spherical Functions on Reductive Groups, Springer-Verlag, 1988.
[13] Y. Guivarc'h, L. Ji, J.C. Taylor, Compactification of Symmetric Spaces, Birkhäuser, 1998.
[14] S. Helgason, Differential Geometry, Lie Groups, and Symmetric Spaces, Academic Press, 1978.
[15] F.B. Knight, A characterization of the Cauchy type, Proc. AMS 55 (1976) 130–135.
[16] F.B. Knight, P.A. Meyer, Une caractérisation de la loi de Cauchy, Probab. Theory Related Fields 34 (1976) 129–134.
[17] G. Warner, Harmonic Analysis in Semi-Simple Lie Groups I, Springer-Verlag, 1972.
[18] R.J. Zimmer, Ergodic Theory and Semi-Simple Groups, Birkhäuser, 1984.